DOCUMENT RESUME

ED 137 346 TM 006 147

AUTHOR Wright, William J.; Sachse, Thomas P.
TITLE Payoffs of Adversary Evaluation.
PUB DATE [Apr 77]
NOTE 20p.; Paper presented at the Annual Meeting of the American Educational Research Association (61st, New York, New York, April 4-8, 1977)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Evaluation Methods; Models; *Program Evaluation; Questionnaires
IDENTIFIERS *Adversary Evaluation

ABSTRACT

The advantages of the adversary evaluation model are argued. Adversary evaluation utilizes a judicial, argumentative approach and promotes the use of multiple types and sources of data. It grants an investigative license to dig out the data to a greater extent than more traditional approaches. Some of the payoffs of the adversary approach are: (1) the critical input of colleagues is considered at the most opportune time--prior to actual investigation; (2) the inclusion of provocative questions which address the real issues of various subgroups is encouraged, resulting in the selection or development of higher quality instruments; (3) the model explicitly encourages the use of qualitative as well as quantitative information; (4) it grants the freedom of investigative license; (5) because polar positions are represented, people feel that issues they believe to be important are not overlooked; and (6) the model provides a built in meta-evaluation. (RC)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED137346

PAYOFFS OF ADVERSARY EVALUATION[1]

William J. Wright
Northwest Regional Educational Laboratory
Thomas P. Sachse
Stanford University

In this paper we will argue the advantages of the adversary evaluation model. Our attempt is to persuade the reader that this approach has substantial merit, and that it ought to be used more frequently than it is. A couple of caveats are in order before we proceed.

In one sense, this symposium is an adversarial examination of the adversary model. We will discuss only the positive aspects of the adversary model. Dr. Popham and Dr. Carlson have the more difficult task of identifying its deficits.

Of course, in any adversarial endeavor it is important to ensure a balance between the teams. The Hawaii 3 on 2 evaluation included two trips to that island paradise. The fact that the co-authors of this paper withdrew from the evaluation prior to the second trip to Hawaii is perhaps sufficient evidence of the relative wisdom of those arguing payoffs versus those arguing deficits.

TM006 147

---

[1] A paper presented as part of a symposium, The Adversary Evaluation Model: A Second Look, Annual Meeting of the American Educational Research Association, New York, April 4-8, 1977.

Second, we base our arguments on a single experience with the model--the evaluation of the Hawaii 3 on 2 program. The literature we reviewed was informative, but essentially we are generalizing from one case--always a dangerous practice. Since that situation has obviously not hindered our colleagues, we will attempt to be as persuasive in proselytizing as they are in condemning.

Peeling the Patina of Professionalism

Many evaluations, like novels, are written in the omniscient third person. The evaluator poses as an entirely objective external party with no vested interest in that which is being evaluated. In truth, we are, none of us, so removed, so above it all, as to lack any bias. Perhaps unconsciously, our pre-dispositions prompt us to plan the evaluation in a certain way. Likely, these biases influence our choice of instruments or data collection techniques, as well.

After an evaluation is concluded and reported, we might receive some gentle chiding from colleagues with a different prospective. Seldom, however, do these differences result in a contentious or public dispute. There are some exceptions, for example, "Sesame Street, Revisited", but these rejoinders seldom receive any great notice, and interest soon dies out.

The adversary approach not only permits, it ensures such disputes, and it virtually guarantees that these arguments will occur early enough to have a positive impact on the evaluation. After a short prelude of congenial fellowship, the two teams can drop all vestiges of gentility and enter the fray with enthusiasm. There are two areas where this tension has real payoff.

Planning the Evaluation

The 3 on 2 program, like Topsy, just grew. There was no predetermined plan for its rapid expansion. Some 3 on 2 classes used the Hawaii English Program (HEP), others did not. Some self contained classes used HEP, but not, proportionately, as many as the 3 on 2 classes. Neither the 3 on 2 program nor HEP was equitably distributed across socio-economic levels. In the face of these, and other confounding variables, it was impossible to contrive an entirely satisfactory sampling plan.

So we argued - almost interminably. The result was, in our view, worth the effort. Each team wanted to ensure that its side was not placed at a disadvantage. Double and triple checking became the order of the day. Because of this attitude, we discovered and corrected certain misinformation about where the 3 on 2 classes were located in time to revise the sampling plan. This new plan avoided a real hazard in that it took into account the hitherto unknown disproportionate allocation of 3 on 2 across SES levels. It wasn't perfect, but the fact that both teams could live with the final sample plan tells us that it was probably as good as one could get.

The whole process reminds one of the terse maxim of the New England carpenter "measure twice - cut once."

Choosing Instruments

The perspectives of the two teams who evaluated the 3 on 2 program concerning the measurement of student achievement were

quite different. This divergence came about, in the main, because of disagreements about fundamental measurement issues.

The chairman of the advocate team was, after all, a leader of the objective referenced testing movement.

Compromise in the choice of instruments was never a real possibility. The arbiters wisely agreed to expend the resources necessary to undertake both norm referenced and criterion referenced testing. (The results were the same in both instances--no significant differences.)

The evaluation of the 3 on 2 program occurred on an extremely tight schedule. The first team meetings did not occur until April of 1976. The final report had to be presented in January of 1977. It was not possible to administer or obtain a measure of the students' aptitudes prior to their entry into a 3 on 2 or a self-contained class. After considerable debate, however, we were able to agree that the administration of the Raven's Progressive Matrices test at the same time as the norm referenced Spring tests would suffice as a covariate.

The development of the questionnaires used in the 3 on 2 evaluation is probably a better illustration of the adversary model's payoff in this area. The developers of these instruments struggled to attain "balance" rather than "neutrality." Provocative, one-sided questions were permissable as long as the instrument, on balance, was acceptable to both teams, and to the project directors. Thus we were able to include in the teacher questionnaire Likert type items with stimulus statements such as:

In the 3 on 2 Program

c. Children do not receive adequate discipline.

e. Children in a 3 on 2 class learn from each other.

f. Children are distracted from learning by the noise and movement in the class.

g. Older children waste time tutoring younger ones.

z. The principals have not provided adequate support for the 3 on 2 teams.

Such questions got right at the heart of the arguments dearest to proponents and opponents of the program, and provided our respondents an opportunity to address themselves to issues which they knew to be real - not bland restatements.

The response rates for our questionnaires were 76.5% for teachers, 82.3% for principals, and 50.8% for parents. No follow up procedures were used. This high return is partly accounted for by the controversial nature of the 3 on 2 program. We speculate, however, that it was also due to the fact that (by means of interviews and a pilot test of the first draft questionnaire) each team ferreted out the issues of concern to that portion of the public (pro or con) it represented, and saw to it that those questions got asked.

Removing the Methodological Manacles

Most evaluations seem to manifest an implicit adherence to Thorndike's belief that "If something exists, it exists in some quantity, and therefore, it can be measured." The aphorism may

in time prove true. As of now, however, we certainly do not possess adequate instruments for the investigation of anything approaching the broad spectrum of questions that ought to be addressed in a comprehensive evaluation. We need to include many information gathering methods in order to even begin to rough in the sketch. Yet, we must also be concerned about objectivity. There are too many uncorraborated, unsubstantiated and unchallenged "observational studies" in our literature already. The adversary evaluation model has real payoff in that it promotes the use of multiple types and sources of data and grants an investigative license to dig out the data to a greater extent than more traditional approaches.

Multiple types and sources of Information

The adversary evaluation model is based upon a judicial system[2] which relies on human testimony - witness and expert. This testimony is subject to cross examination and the corroboration or refutation of other witnesses and experts. The adversary model thus explicitly encourages the use of non-quantitative information. Quantitative data is most assuredly permissable, but it is not per se more valuable than other information.

One cannot argue that the adversary evaluation model is unique in this regard. Many non-adversarial evaluations have included a wide spectrum of information gathering mechanisms. The adversary model is distinctive, however, in that, by its very nature, it fosters the use of multiple sources and types of data.

---

2 The adversary evaluation model is not necessarily judicial, of course. It can be structured as a debate or an arbitration proceeding. The nature of the evidence permissable is the same in any case.

Since each side is seeking whatever information is most supportive of its position, a wider diversity of data may be collected in an adversary evaluation. In the 3 on 2 evaluation the teams used the following approaches:

- Norm referenced measures of student achievement
- Criterion referenced measures of student achievement
- Interviews with

    Legislators
    Members of the State Board of Education
    Staff of the Department of Education
    Principals
    Teachers
    Parents
    Students

- Student attitude Inventories
- Questionnaires

    Principals
    Teachers
    Parents

- Analysis of Teacher Absenteeism Records
- Analysis of Cost and Budget Information
- Classroom Observation

    Structured
    Unstructured

- Analysis of Historical Documents
- Public Meetings

This panoply of information was not part of our original plans, rather it resulted from the sometimes almost desperate search for evidence favorable to one's cause. Indeed it was somewhat amusing to observe hard nosed testing specialists willingly praise the virtues of and press the necessity for classroom observation and description. It was rather like watching Hemingway running amok--with a butterfly net.

Other evaluation methods use both quantitative and qualitative data, of course. The important difference between the adversary model and other approaches rests in the sphere of objectivity. Objectivity has been defined in terms of an agreement among competent observers. When the advocate and adversary agree, one can rest fairly certain that the phonomenon is real.

Viewed from a different perspective, meta-evaluation is built into the adversary model. You don't dare crawl too far out on any evidential limb. The opposing side is observing your perch--saw at the ready.

Digging out the Data

The adversary evaluation model also pays off in that it grants evaluators investigative license. They become detectives with a client. There is no need to bend over backwards to avoid offending.

Those who were interviewed in the 3 on 2 evaluation were usually interviewed simultaneously by one person from each team. These "witnesses" understood, and frankly seemed to enjoy the situation. Tough questions were expected. The evaluators had the right, indeed the obligation, to ask for supporting details whenever a statement about the program was made. We also had the right to search for evidence that contradicted the testimony of even the most influential.

Some examples of what this investigative license can lead to are in order. We interviewed district superintendents, principals and teachers about such issues as how the decision was made to install the 3 on 2 program in a given school. The answers were

often contradictory. The principal would state he left the decision up to his teachers, for example, but the teachers would insist that they had no voice in the matter. Enough confusion was evidenced in our preliminary interviews for us to address this and similar issues in the questionnaires. As a consequence we learned that considerable confusion existed about who was responsible for decisions about implementing the 3 on 2 program. This evidence on the lack of clear leadership might not have been as vigorously pursued or as frankly reported had another evaluation approach been used.

A more pointed example of this payoff occurred in the examination of the costs of the 3 on 2 program. We were originally led to believe that the cost of the program was about 6 million dollars a year. Indeed that figure had been widely circulated and accepted as accurate by a number of those with whom we talked. By some calculations the 6 million dollar figure seemed perfectly accurate. Cost was a major issue in this evaluation, however, and both teams wanted to learn as much as they could about the expenditures required for the 3 on 2 program. Naturally the advocate team hoped to find lower costs or cost savings; whereas the adversary team wanted evidence that the program was inordinately expensive. So by digging into various budget reports and asking probing questions of several individuals, the two evaluators (one from each team) assigned to cost data determined that the program actually cost in excess of 9 million dollars per year. Without the legitimacy of a cause to represent, we question whether they would have been as persistent or as insistent in securing accurate information.

Pleasing The People

Evaluation is often perceived more as beast than beauty by those who call upon its services. Its unpopularity stems from many factors, but perhaps primary among them is the insensitivity with which traditional evaluation approaches are often applied. Any evaluation has many audiences, each of which has its own questions of interest. Only a small portion of these questions usually get asked. Moreover, most evaluations ignore the vast bulk of those who want an opportunity to express their opinion or who at least want to be informed about the progress and results of the evaluation.

The adversary approach promotes the use of multiple types and sources of data. It also has a distinctive and public means of reporting. The 3 on 2 program featured, for example, an hour long televised presentation of a debate between representatives of each team. Because the people affected by the decisions to be made have more of an opportunity to be involved in the decision making process, we believe the adversary approach elicits their support rather than their enmity. To check this perception we decided to use a simple questionnaire.

The Questionnaire Study

The questionnaire which we used was developed jointly by Blaine R. Worthen and William J. Wright. It was intended for those most immediately involved in making or implementing

decisions about the future of the 3 on 2 program. It concerned only the value and utility of the adversary approach. A copy of that questionnaire with the number of individuals who responded to each response option filled in is appended to this paper.

The questionnaire was sent to 34 individuals including each member of the Hawaii Board of Education, each district and deputy district superintendent, and the principal staff members of the Department of Education.

One of our respondents was too ill to oblige us. Of the remaining 33, 27 or 82 percent, returned the questionnaire, some anonymously. We will report on those items most directly related to client satisfaction in this paper. Blaine Worthen will address the remaining issues in his presentation.

The majority of those responding to the questionnaire (78 percent) believed that "the advocate-adversary approach more clearly focuses on issues relevant to the decisions to be made." Only <u>one</u> respondent felt that "the advocate-adversary approach tends to confuse rather than illuminate the issues."[3]

Among the other characteristics identified by our respondents as descriptive of the adversary evaluation approach were the following:

---

[3] Since the directions stated "check <u>any</u> that apply," and responses were not polar, the sum of those who chose a particular set of statements will not equal 100% except by chance.

- "The advocate-adversary approach provides more comprehensive information about the program being evaluated." (55 percent)
- "The advocate-adversary approach results in a more balanced presentation--important issues do not get overlooked." (78 percent)

Few of our respondents felt that the following statements were true:

- "The advocate-adversary approach depends too much on the personality characteristics of the spokespersons for the two teams." (33 percent)
- "The advocate-adversary approach focuses too much attention on the evaluation method rather than on the program being evaluated." (7 percent)
- "The advocate-adversary approach does not provide decision makers with the evidence they need to make a choice." (7 percent)

The most impressive evidence of the payoff of the adversary approach with respect to client satisfaction came in response to the question "Would you recommend the use of the advocate-adversary evaluation approach to others?" Of the 25 individuals who answered the question, 10 (or 40 percent) responded "definitely yes"; 14 (or 56 percent) responded "probably yes"; 1 respondent was "uncertain". There were no negative responses to this question!

Conclusion

In this section of the paper we will review briefly the payoffs of the adversary model. We will then attempt to look

beyond the 3 on 2 experience and suggest some variations to the adversary model that might be interesting to try out. Finally, we will look at what may be the biggest payoff of all.

Summing up the Payoffs

The payoffs of the adversary approach are many. The listing which follows is intended as a quick summary.

- Planning

  The adversary model has real benefits in that, as the two teams plan the evaluation, debate about critical issues such as problem definition, sampling plans and the like occur. Thus the critical input of colleagues is considered at the most opportune moment--prior to actual investigation.

- Instrumentation

  Because of the adversarial nature of this approach, the instruments selected or developed are likely to be of higher quality. In the development of questionnaires and interview schedules the focus is on balance not blandness. The inclusion of provocative questions which address the real issues of various sub groups is thus encouraged.

- Multiple Types and Sources of Information

  The adversary model explicitly encourages the use of qualitative as well as quantitative information. It compels the evaluators to search vigorously for any credible evidence supportive of the position they champion. Yet, because of the constant presence of

evaluators who represent the contrary position, objectivity, in the sense of agreement among competent observers, is at a premium. When advocate and adversary report observing the same phenomenon, one can have confidence in its reality.

- Investigative License

  The advocate or adversary evaluator represents a particular cause. In this situation he or she has more freedom to challenge testimony and to dig into records, reports, and other sources of information seeking corroborative or contradictory evidence. Sometimes, when you seek to ferret out the truth, you find it.

- Client Satisfaction

  Because the advocate and adversary evaluators represent the polar positions within some larger community, people feel that the issues they believe to be important are not overlooked. The public nature of the reporting mechanism also involves people. They don't feel as excluded from the decision making process. In a study conducted among our principal client group, all but one respondent to a question concerning whether they would recommend the adversary model to others reported that he or she would probably or definitely recommend the adversary evaluation model to others.

- Built in Meta Evaluation

  Running through all the payoffs is the notion of a built in meta evaluation with a distinct advantage. Instead of a carping critic of the Monday morning quarterback variety, the meta evaluator is an opponent, who is there when needed, and who has as large an interest in the ultimate credibility of the evaluation as you do.

Looking Beyond the 3 on 2 Experience

The approach used in connection with the 3 on 2 program is but one type of adversary evaluation model. Owens and Hiscox have described other versions in their paper. We would like to suggest that the investigation of other judicial models may provide some interesting variations. In some European courtrooms, for example, the judge is not an impartial monitor, but is charged with the responsibility of seeking the truth of the matter. The judge is free to, and often does, cross question witnesses as vigorously as either of the attorneys. In certain systems all members of the bar have an equal responsibility to find the truth. The defense, prosecution, and judge work toward the same end, but from different perspectives. Facts are established and clearly distinguished from opinions and recollections which are disputable. Perhaps something akin to this approach might be of value in education.

Yet we must be careful to learn what we can from our own judicial system before we move on too quickly to other models.

Consider the fact that those who pass our laws, those who enforce the laws, those who dispute the legal issues of the day and those who sit in judgement of these disputes all tend to be members of the same professional society. Consider the benefits of such a system. There is much to be learned indeed!

The Big Payoff

The big payoff of the adversary evaluation model is that it forces us, as a profession, to explore alternative means of obtaining and reporting information. It is not the answer, but it is one more arrow in our arsenal, and we need all the weapons we can get.

REFERENCES

Cook, T. D.; Appleton, H.; Connor, R. F.; Shaffer, A.; Tamkin, G.; Webber, S. J.; Sesame Street, Revisited, Russell Sage Foundation, New York, 1975, 410 pages.

3 on 2 Evaluation Report, Volume 1, Technical Report, Portland, Oregon: Northwest Regional Educational Laboratory, Assessment Program, January 1977.

3 on 2 Evaluation Report, Volume 2, Team Interpretations and Recommendations, Portland, Oregon: Northwest Regional Educational Laboratory, Assessment Program, January 1977.

3 on 2 Evaluation Report, Volume 3, Executive Summary, Portland, Oregon: Northwest Regional Educational Laboratory, Assessment Program, January 1977.

Name_____

1. Now that you have had an opportunity to read the written reports and hear the oral reports of the 3 on 2 evaluation, do you feel that the advocate-adversary approach produces information worth the cost of having two teams involved? (check ONE)

15% [4] The information produced by advocate-adversary evaluation was worth more than the evaluation's cost

78% [21] The information produced by the advocate-adversary evaluation was worth the evaluation's cost

7% [2] The information produced by the advocate-adversary evaluation was not worth the evaluation's cost

2. Compared to other methods of evaluation, which of the following statements about the advocate-adversary approach to evaluation do you think are true? (Check ANY that apply)

78% [21] The advocate-adversary approach more clearly focuses on issues relevant to the decisions that are to be made

4% [1] The advocate-adversary approach tends to confuse rather than illuminate the issues

33% [9] The advocate-adversary approach depends too much on the personality characteristics of the spokespersons for the two teams

56% [15] The advocate-adversary approach provides more comprehensive information about the program being evaluated

78% [21] The advocate-adversary approach results in a more balanced presentation--important sub-issues do not get overlooked

7% [2] The advocate-adversary approach focuses too much attention on the evaluation method rather than on the program being evaluated

44% [12] The advocate-adversary approach presents more clearly defined decision alternatives and the justification for each

7% [2] The advocate-adversary approach does not provide decision makers with the evidence they need to make a choice

3. In which of the following circumstances do you feel it would be appropriate to use an advocate-adversary evaluation? (check ANY that apply)

78% [21] When decisions must be made about whether to continue or terminate a program

15% [4] When decisions must be made about how to improve a program

56% [15] When the evaluation is conducted by external evaluators (persons outside of the program being evaluated)

15% [4] When the evaluation is conducted by internal evaluators (persons employed by the program being evaluated)

74% [20] When the program being evaluated is large and affects a large number of people

93% [25] When the program being evaluated is controversial and many people hold strong positive and negative opinions of it

63% [17] When there are many different audiences for the evaluation report

30% [8] When it is important to have an interesting non-written method of reporting the evaluation results

Other (please explain)_____

_____

_____

_____

4. Would you recommend the use of the advocate-adversary evaluation approach to others?

Definitely Yes	Probably Yes	Uncertain	Probably No	Definitely No
[10]	[14]	[1]	[ ]	[ ]
37%	52%	4%		

5. What do you feel are the unique advantages (if any) of the advocate-adversary approach to evaluation?

6. What do you feel are the unique disadvantages (if any) of the advocate-adversary approach to evaluation?